1980-03

The use of parametric cost estimating relationships as they pertain to aircraft airframes: a new perspective

Bennett, Bruce Robert

Monterey, California. Naval Postgraduate School

http://hdl.handle.net/10945/19098

THE USE OF PARAMETRIC COST ESTIMATING
RELATIONSHIPS AS THEY
PERTAIN TO AIRCRAFT AIRFRAMES; A NEW
PERSPECTIVE

Bruce Robert Bennett

NAVAL POSTGRADUATE SCHOOL
Monterey, California

THESIS

The Use of Parametric Cost Estimating Relationships as They Pertain to Aircraft Airframes; A New Perspective

by

Bruce Robert Bennett

March 1980

Thesis Advisor: M. G. Sovereign

Approved for Public Release; Distribution Unlimited

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

REPORT DOCUMENTATION PAGE		READ INSTRUCTIONS BEFORE COMPLETING FORM
1. REPORT NUMBER	2. GOVT ACCESSION NO.	3. RECIPIENT'S CATALOG NUMBER
4. TITLE (and Subtitle) The Use of Parametric Cost Estimating Relationships as They Pertain to Aircraft Airframes; A New Perspective		5. TYPE OF REPORT & PERIOD COVERED Master's Thesis; March 1980
		6. PERFORMING ORG. REPORT NUMBER
7. AUTHOR(s) Bruce Robert Bennett		8. CONTRACT OR GRANT NUMBER(s)
9. PERFORMING ORGANIZATION NAME AND ADDRESS Naval Postgraduate School Monterey, California 93940		10. PROGRAM ELEMENT, PROJECT, TASK AREA & WORK UNIT NUMBERS
11. CONTROLLING OFFICE NAME AND ADDRESS Naval Postgraduate School Monterey, California 93940		12. REPORT DATE March 1980
		13. NUMBER OF PAGES 46
14. MONITORING AGENCY NAME & ADDRESS(if different from Controlling Office) Naval Postgraduate School Monterey, California 93940		15. SECURITY CLASS. (of this report) Unclassified
		15a. DECLASSIFICATION/DOWNGRADING SCHEDULE

16. DISTRIBUTION STATEMENT (of this Report)

Approved for Public Release; Distribution Unlimited

17. DISTRIBUTION STATEMENT (of the abstract entered in Block 20, if different from Report)

18. SUPPLEMENTARY NOTES

19. KEY WORDS (Continue on reverse side if necessary and identify by block number)

Cost Estimating Relationships, Aircraft Airframe Costs, Mahalanobis Distance

20. ABSTRACT (Continue on reverse side if necessary and identify by block number)

The purpose of this thesis was to review cost estimating relationships that have been developed and used for aircraft airframe costs, to identify existing problems, and where appropriate, to suggest alternatives for the future application of cost estimating relationships to aircraft airframes. Mahalanobis distance was explored as a means of complementing the more traditional statistical measures for regression analysis. This study supports the conclusion that cost estimating relationships should be developed for a specific system to be estimated, and that Mahalanobis distance is a potentially effective

DD FORM 1 JAN 73 1473 EDITION OF 1 NOV 65 IS OBSOLETE
(Page 1) S/N 0102-014-6601

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

tool by which the analyst may address the important issue of analogy between the data base and the proposed system.

The Use of Parametric Cost Estimating Relationships as They Pertain to Aircraft Airframes; A New Perspective

by

Bruce Robert Bennett
Lieutenant Commander, United States Navy
B.S., University of Wisconsin, 1970

Submitted in partial fulfillment of the
requirements for the degree of

MASTER OF SCIENCE IN OPERATIONS RESEARCH

from the

NAVAL POSTGRADUATE SCHOOL

March 1980

ABSTRACT

The purpose of this thesis was to review cost estimating relationships that have been developed and used for aircraft airframe costs, to identify existing problems, and where appropriate, to suggest alternatives for the future application of cost estimating relationships to aircraft airframes. Mahalanobis distance was explored as a means of complementing the more traditional statistical measures for regression analysis. This study supports the conclusion that cost estimating relationships should be developed for a specific system to be estimated, and that Mahalanobis distance is a potentially effective tool by which the analyst may address the important issue of analogy between the data base and the proposed system.

TABLE OF CONTENTS

LIST OF TABLES

I. INTRODUCTION

An independent parametric cost estimate is defined in Reference 1 as an estimate which predicts cost by means of explanatory variables such as performance characteristics, physical characteristics, and characteristics relevent to the development process, as derived from experience on logically related systems. It is a means to an end. Decisions that inevitably have to be made are based in part on what has happened in the past, and in part, on what is expected to happen in the future.

One of several areas within DOD where uncertainty about the future hinders the decision-making process is in the acquisition of major weapons systems. The need to determine a "priori," the cost impact of such a decision, is important from a budgeting point of view, and with the increased fiscal constraints, the cost impact of a decision can be as significant as the performance characteristics of the system desired.

Typically, the choice among systems is based on trade-offs between various performance parameters in attempting to determine which system will best fulfill the mission requirements. In the past, cost was not always a major consideration in defining the requirements. However, given the requirements, every effort was made to procure them at the best possible cost to the government.

In an attempt to save more money in the long run, and operate within tighter budgets, DOD instruction 5000.1 was issued. It defines specific design to cost policies and upgrades cost to a principle design parameter. Cost must now be considered during requirements formulation in determining which system provides the best value in fulfilling mission needs.

This situation is recognized at all levels within DOD as evidenced by a great number of policy directives concerning the problems with cost overruns and the need to improve cost estimating proceedures. In 1971, the Deputy Secretary of Defense directed each of the Service Secretaries to: 1) improve their capability to perform independent parametric cost estimates; 2) utilize their capability at all key decision points in the acquisition process, and 3) insure that the results of the analysis are made available to the Defense System Acquisition Review Council (DSARC) at each DOD program milestone.

In a report to Congress one year later, the General Accounting Office (GAO) recommended in part that "DOD develop and implement guidance for consistent and effective cost estimating proceedures and practices, particularly with regard to . . . an effective independent review of cost estimates." As a result of this and other impetus, considerable effort has been expended in attempting to develop suitable cost estimating relationships (CER). A CER is a mathematical expression that determines cost as a function of various system characteristics. Either directly or through proxy, these system characteristics determine the value of the explanatory or independent variables that comprise the functional form. "The construction and use of CERs form the foundation for making independent parametric cost estimates."[1]

There are several reasons why CERs have been and will continue to be important in the acquisition process. Early in the process when many alternative designs are contemplated, a CER based on readily available performance characteristics (explanatory variables) allows the decision

---

[1]Miller, Bruce M. and Sovereign, Micheal G., <u>Parametric Cost Estimating with Application to Sonar Technology</u>, p. 2, Naval Postgraduate School, NPS 55207309lA, September 1973.

maker to evaluate the cost impact of the various designs (or changes thereof) and make trade-offs accordingly. To attempt this type of analysis with other than a CER would be both cost and time prohibitive.

As requirements become more defined and other estimates are made available a CER can be used to verify their potential accuracy. For example, after receipt of several contractor proposals for a specific weapons system, CERs developed for individual cost elements may well indicate areas where the contractor may have "padded" his estimate, or perhaps misinterpreted the specification requirements. This is especially true when solicitation specifications are performance oriented, allowing the contractor more latitude in design and thus significant differences among the various proposals. After acquisition, and well into the production phase of a weapons system, the potential use of a CER still exists. Major changes in design (either contractor or government initiated) may be extensive enough to warrant the use of a CER as an initial determination of cost, or to verify a more detailed engineering estimate.

Recognizing the need for and usefulness of a parametric cost estimating relationship is the easy part. Developing a reliable CER is difficult at best. There are many problems the analyst must overcome in achieving this end. Identifying and collecting the data is the first and most difficult obstacle. The availability of cost information for a number of previously acquired "similar" systems is important. Application of CERs to the aircraft acquisition process has received considerable attention, in part because a reasonably large number of aircraft have been procured since 1950 for which cost information is available.

Several techniques/methods for determining an appropriate CER have been tried and are continually being massaged. This thesis effort is an attempt to summarize these methods as they relate to aircraft airframe costs, to identify trends and limitations, and to address the appropriateness of a shift in direction to enhance the future usefulness of parametric cost estimating techniques.

II. BACKGROUND AND TRENDS IN COST ESTIMATING RELATIONSHIPS

The development of a cost estimating relationship (CER) is dependent upon the existence of historical information. The ultimate quality of the CER (its ability to accurately predict costs) can be no better than the data upon which the CER was based.

DOD recognized the need for and the difficulty of data collection in the early 1960s. At this time the only information available was that provided under government contract, either as a part of the initial proposal or, as in the case of cost-type contracts, as part of the billing and audit processes. Information could, and still can be, obtained directly from the manufacturer if they choose to provide it, but as with the case of DOD secured information, it was both sporadic and inconsistent. It was inconsistent in the sense that there were no standards by which manufacturers were required to accumulate and report costs.

In an attempt to correct these inadequacies, the Contractor Information Report Program (CIR) was implemented in 1966. It was designed to collect specific cost related information on major contracts for aircraft, missiles, and space programs. It has subsequently been enlarged to include other programs and is now referred to as the Contractor Cost Data Reporting System (CCDR).

In addition, the initiative was taken to standardize proceedures by which costs would be accumulated and reported. This was accomplished by the Cost Accounting Standards Board and based on establishing consistency of accounting practices among government contractors. Admittedly, the motive of this action was to enhance the DOD contracting

personnel's ability to evaluate proposals and better determine allocability and allowability of costs, but an obvious additional benefit was to create some consistency in the data base.

Each major airframe manufacturer has developed their own data base and corresponding models. They are used quite extensively by these manufacturers in their design selection process and in the preparation of proposals. Because of the selective nature of the sample from which they are derived, their use is considered limited, but the techniques employed to develop them will be discussed later.

On an industry-wide basis, DOD must be considered the ultimate repository of the most accurate and current military aircraft airframe cost information. It would not be possible for any organization outside of DOD to replicate this data base, primarily because of the proprietary basis upon which most of the information was received.

Mainly in support of Air Force sponsored research efforts, through the years the Rand Corporation has organized and updated the DOD data base for airframe costs, identifying the deficiencies and correcting them where possible. For each of the forty-three (43) aircraft in the existing data base, costs are provided for seven (7) different categories. The two pre-production nonrecurring cost categories include flight test costs and development support costs. Cumulative totals for the remaining five (5) production related categories include engineering hours, tooling hours, recurring manufacturing labor hours, manufacturing material dollars, and quality control hours. The cumulative totals that are provided are for production quantities of 25, 50, 100, and 200 units and are based on a fitted cost versus quantity curve which was extrapolated if actual production quantities were less than 200 units.

In using this data (as with any other data base) the analyst must be familiar with its derivation and aware of its deficiencies. As implied earlier, many of the deficiencies that exist are a result of compiling data submitted by many contractors utilizing different accounting practices. The overhead accounts are an example of where this might occur. Part of the differences in cost may be attributed to a difference in the allocation base. Another example of a possible source of error is tooling costs that occur during the production process and should be recorded as a nonrecurring cost, but are often included in the production oriented recurring costs. The need for recognizing these sorts of problems in developing a CER will be explored in more detail in section III of this paper in the context of adjusting raw data.

Many organizations have developed cost models and several techniques/methodologies have been employed. By reviewing some of these methods, the reader should gain an understanding of where the emphasis has been placed and what trends have been established.

The Rand Corporation has used the data base discussed earlier in this section. Regardless of mission profile or type, all aircraft in the sample were used, with the exception that for each revision of their present model some older aircraft were deleted and the more recent aircraft added. This was done for several reasons. The cost information for older aircraft was less reliable than for later aircraft, and the development and production experience of these earlier aircraft were not considered an appropriate indicator of the future. The current Rand model, DAPCA III, is based on a sample of twenty-five (25) aircraft, all of which have a first flight date of 1952 or later.

In selecting the explanatory variables for their CER, Rand used the following guidelines: "1) They must be quantifiable early in the

design phase. 2) Certain preconceived relationships to cost must be supported by the CER. 3) They must be statistically significant."[2] The first requirement implies that it is useless to have a CER to estimate future cost if detailed information is required in order to determine an appropriate value for the explanatory variable. The time of first flight is an example of an explanatory variable that is hard to quantify early in the decision process when actual performance characteristics have yet to be definitized. The second requirement is an attempt to avoid spurious correlation, and the third requirement insures that the explanatory variables are in fact contributing to explaining the variability in the data.

A log-linear functional form has traditionally been used by Rand because of the implied diminishing marginal returns when coefficients are less than 1.0. In this context, coefficient values greater than 1.0 became grounds for questioning the merit of the particular explanatory variable.

Utilizing this functional form, a regression analysis was done in each of the seven (7) cost categories for many combinations of as many as twenty (20) different explanatory variables. The coefficient of determination (R^2) was used as a first cut to determine the better CERs. The guidelines for explanatory variables having been employed, the causal relationships to cost could be supported. The final test was how well the CER performed in predicting the cost of the more recent aircraft. In all cost categories, the "optimal" CER used weight and speed as the

---

[2] Large, J. P., Campbell, H. G., Cater, D., Parametric Equations for Estimating Aircraft Airframe Costs, p. 4, Rand Corporation Report R-1693-PA&E, May 1975.

explanatory variables. There were two exceptions to this: manufacturing labor and manufacturing materials use an optional third explanatory variable that is related to time.

Since DAPCA III was published in 1976 (Table One, compiled from Ref. 2), the Rand Corporation has pursued the use of other explanatory variables that were felt would be better predictors than just weight and speed. One reason for this was the result of the work of Timson and Tihansky (Ref. 17) which criticized the size of the prediction interval for the DAPCA III CERs.

In the pursuit of better predictors of cost, two of the most promising areas were defining a measure of technological trends and identifying reasonably quantifiable program related explanatory variables. Reference 15 is a detailed report on the most recent work in quantifying technological advance in aircraft. Using explanatory variables that measure aircraft performance (e.g., specific power, range, sustained load factor) a relationship was developed using multiple regression that determines time of first flight of a particular aircraft as a function of these performance characteristics. The obvious next step was to use this measure of technological advance to help explain differences in cost. This was attempted and the results are summarized in Ref. 5. It met with limited success, in part, due to the correlation between the time of first flight and any performance oriented explanatory variable that was used in the CER.

The most recent model developed by the Planning Research Corporation (PRC), which was published in 1967, is quite different from the Rand approach. It was designed to be used after a contractor has been chosen and a production schedule has been defined. The data base consists of

TABLE ONE

SELECTED CERs FROM THE RAND CORPORATION MODEL (DAPCA III)

$$E = 20.032 \cdot W^{0.6636} \cdot S^{0.9871} \cdot 200^{-(b+1)} \cdot Q^{b+1} \cdot 10^{-6}$$

$$T = 522.39 \cdot W^{0.6214} \cdot S^{0.5323} \cdot 200^{-(b+1)} \cdot Q^{b+1} \cdot 10^{-6}$$

$$ML_{NR} = 0.62597 \cdot W^{0.6883} \cdot S^{1.2109} \cdot 10^{-6}$$

$$ML_{R} = 1188.5 \cdot W^{0.8306} \cdot S^{0.5464} \cdot T^{-0.4711} \cdot 200^{-(b+1)} \cdot 10^{-6}$$

$$ML_{R} = 581.55 \cdot W^{0.7830} \cdot S^{0.4297} \cdot 200^{-(b+1)} \cdot Q^{b+1} \cdot 10^{-6}$$

$$MM_{R} = 191.85 \cdot W^{0.8600} \; s \; S^{0.8126} \cdot 200^{-(b+1)} \cdot Q^{b+1} \cdot 10^{-6}$$

$$FT = 153.25 \cdot W^{0.7095} \cdot S^{0.5856} \; Q_{FT}^{0.7160} \cdot DV^{-1.5570} \cdot 10^{-6}$$

Where:

E = total engineering hrs (millions)

T = total tooling hrs (millions)

ML_{NR} = nonrecurring manufacturing labor hours (millions)

ML_{R} = recurring manufacturing labor hours (millions), with or without time variable

MM_{R} = recurring manufacturing materials (millions of 1975 dollars)

FT = flight-test costs (millions of 1975 dollars)

W = airframe unit weight (lb)

S = maximum speed at best altitude (kn)

b = determined from cumulative average slope of anticipated learning

Q = airframe quantity

Q_{FT} = number of flight test aircraft

DV = dummy variable (2 = cargo, 1 = all other)

twenty-nine (29) aircraft with first flight dates that range from 1945 to 1958. Only four (4) cost categories are used, and all information is given in dollars except for manufacturing labor. The four cost categories are: 1) Nonrecurring tooling and engineering dollars. 2) Recurring tooling and engineering dollars. 3) Manufacturing labor hours (includes quality control). 4) Manufacturing material dollars. Two of several possible reasons for this choice of categories include: They are sufficient to fulfill the intent of the CER; and, more detailed cost information is not available for the older aircraft in the sample.

Details as to the basis for developing the CERs used in the PRC model are not completely available. A log-linear functional form is used, and the emphasis on the choice of explanatory variables would appear to be their logical importance relative to cost rather than their statistical significance. The CER for manufacturing material uses speed, a time factor, unit weight, and delivery rate as explanatory variables with speed being the only variable that is significant at the 90% level. As expected, with this type of emphasis on the choice of explanatory variables, a different CER is developed for each cost category.

The remaining model to be discussed, developed by J. Watson Noah Associates, uses yet another approach. The most extensive data base of the three models is used by Noah. It includes thirty-five (35) aircraft with first flight dates that range from 1947 to 1974. In the initial model, the cost information is divided into only two categories --recurring and nonrecurring. In the revised model published in 1977 (Table Two), the categories were redefined as development and production costs (to include all tooling costs). Although the initial model used an arithmetic functional form, the revised model used the log-linear form as used by both the Rand and PRC models.

TABLE TWO

CERs FROM THE J. WATSON NOAH ASSOCIATES MODEL

$$\ln D = -13.013214 + .606684 \ln W + .602425 \ln S - .791948 \ln GW + .877138 \ln F + 1.755809 \ln TI$$

$$\ln P = -8.246325 + .395885 \ln W + .166260 \ln S + .506351 \ln F$$

where,

D = design costs in millions of 1975 dollars

W = airframe unit weight (lb)

S = maximum speed at best altitude (kn)

GW = gross weight (lb)

F = maximum thrust (lb)

TI = technology index

P = cumulative average production cost for quantity 100 in 1975 dollars

Note: Multiply Design Costs by:

1.775393 for bomber aircraft

2.185003 for major technology advance

Multiply Production Costs by:

.727219 for cargo aircraft

1.199087 for bomber aircraft

1.389824 for major technology advance

As with the PRC model, information about the choice of explanatory variables is unclear. It would appear that the emphasis was again placed on logical rather than statistical significance as evidenced by the CER for design costs which contains as two of its explanatory variables, airframe unit weight and gross weight, which are highly correlated. Noah's model also differs from the other two in that it contains an index of technological advance and a judgmental complexity factor. The index of technological advance is basically just a value that is assigned according to the sequential ordering of first flight dates of all aircraft manufactured, whether used in the sample or not. The judgmental complexity factor is based on the ability to single out major differences from earlier aircraft as opposed to what would be considered a normal trend in design or program changes. The CERs for both development and production costs are sensitive to this complexity factor, therefore a proper choice is required to achieve a reasonably accurate estimate.

It is apparent from reviewing these three models that the methods used to determine a CER, and the CERs themselves, are as varied as the number of attempts to develop them. A closer look at the problems and limitations of these CERs and methodologies is required before an attempt to improve and/or consolidate proceedures can be made.

III. LIMITATIONS OF, AND PROBLEMS WITH EXISTING CERs

There are obvious limitations to any cost estimating relationship. Even with perfect historical information, regression theory states that the width of the prediction interval about an estimate increases as the system being considered extends beyond the limits of the data base. The multi-dimensional form of the prediction interval equation is given in Ref. 16 as: $PI = C \pm (t_{1-\frac{\alpha}{2}})\ SE \sqrt{1 + E'\ (X'X)^{-1}\ E}$

where,

C = point estimate of the cost of the system predicted from the regression

$t_{1-\frac{\alpha}{2}}$ = t statistic (constant for a particular CER with α specified)

SE = standard error of the regression model

E = vector of proposed system explanatory variable values, the first element of which is a one (1) to represent the constant term of the regression

X = matrix, each column of which is the value of explanatory variables of a system in the data base. The first column is all ones (1's) and represents the constant term.

Considering for the moment that all other terms are constant, the width of the prediction interval varies according to $E'\ (X'X)^{-1}\ E$. When E equals the column means of X, this expression reduces to $\frac{1}{n}$, where n is the number of systems in the data base. The expression under the radical therefore becomes $1 + \frac{1}{n}$ which can be written as $\frac{n+1}{n}$. This is consistent with the one dimensional form of the prediction where the term under the radical is; $\frac{n+1}{n} + \frac{(E - \bar{X})^2}{(X_1 - \bar{X})^2}$ and reduces to $\frac{n+1}{n}$ when $E = \bar{X}$.

It is interesting to note that the value of the E vector (proposed system characteristics) is not affected by the corresponding values of the X matrix (data base system characteristics). Also, the expression X'X, if adjusted for column means and sample size would result in a covariance matrix for the explanatory variable values of the data base. A technique which incorporates these concepts will be discussed in Section V.

The accuracy of the estimate (i.e., the width of the prediction interval) can only get worse if additional errors are introduced as a result of inconsistencies in available data. These limitations are generally recognized and accepted by the analyst. There are other limitations and problems with CERs, the proposed solutions to which analysts do not readily agree. These problems invariably arise as a result of the shift in emphasis between statistical considerations and judgmental factors, and can usually be shown to account for differences in the existing models. The implication here is that the non-quantifiable aspects of developing and applying a CER result in the use of different techniques which cannot be objectively evaluated. To explore some instances which give rise to these differences is necessary to acquire a better appreciation of the problems that exist.

It may be easy to support a causal relationship between an explanatory variable and cost, but in the resulting CER the coefficient of this variable may be statistically insignificant. Retaining this variable in the CER may give a more logically oriented CER, but if the variable does not contribute appreciably to explaining historical variations in cost, there is no reason to believe that it will be an adequate estimate of change in future explanation of variations in cost.

(In Section II it was shown that Rand chose to disregard the variable, and PRC and Noah chose to retain it.)

A prerequisite for inclusion of an explanatory variable should be the perceived existence of a causal relationship to cost so it is unlikely that a CER with a statistically significant variable with no apparent causal relationship to cost will exist. What can happen, however, is the existence of a statistically significant variable with obvious effects on cost, but extremely difficult to quantify. This is the case with Noah's complexity factor. It is hard to determine if a system will be significantly "different" from historical trends, yet a correct decision is critical to the accuracy of the estimate of cost using this CER. These situations create dilemmas for both the analyst and the user.

Multicollinearity is another problem. It arises when two or more explanatory variables (or combinations thereof) are highly correlated with each other. When multicollinearity exists, interpretations of the coefficients becomes difficult. The coefficient of the first of two correlated variables is a measure of the change in cost for a given change in this variable, all other things considered equal, but due to the collinearity, the values of the second variable also will change. "Because multicollinearity is dependent upon the sample of observations, little can be done to resolve it unless more information about the process in question is available."[3] An understanding and careful choice of explanatory variables is necessary to deal with this problem of multicollinearity.

---

[3]Pindyck, R. S. and Rubinfeld, D. C., Econometric Models and Economic Forecasts, p. 68, McGraw-Hill, Inc., 1976.

Selection of the systems to be used in the data base requires a trade-off between similarities with the proposed system versus sample size. Noah's use of all available aircraft emphasizes sample size, but older aircraft may not accurately reflect more recent trends in production and manufacturing processes or requirements. A more selective homogeneous sample choice may be criticized because typically the size of the sample will become statistically small. Part of the reason for this criticism is evident from the confidence interval formula previously introduced. The t statistic for a fixed α is a function of the sample size n. For small n, the t statistic, and hence the confidence interval, becomes larger. However, this effect is small compared to others.

From a broader perspective, the problems with existing CERs can be attributed to the lack of definition of two basic concepts. The first is the fact that there is not a universally accepted method of measuring how well the data base and the proposed system relate. This relation can be thought of as an analogy between the systems in the data base and the systems to be estimated. The second concept is the tendency to seek or use one "overall best" CER for all applications.

Concerning the first concept, the coefficient of determination (R^2) has been used traditionally as an indicator of how well the estimating relationship (determined by the regression) fits the data. It is a measure of the proportion of total variance of the independent variable from its mean value that is explained by the estimating relationship. Because it is a ratio of variances (i.e., the explained variance divided by the total variance) it is a relative measure that can be used to compare different estimating relationships according to their ability to explain the variances of the dependent variable, which for a CER is cost.

There are two weaknesses associated with the use of R^2. As with any numerical proceedure, it lacks the ability to identify the existence of a causal relationship between independent and dependent variables. It is realized that this problem only can be addressed by the analyst in his selection of explanatory variables. It is presented here only for completeness. Of concern in the use of R^2 is the fact that its value is completely determined by the data base. The nature of the system to be estimated has no effect on its value. In essence, it lacks a measure of analogy that the analyst should use to determine an appropriate data base given the characteristics of the system to be estimated. It is not presumed that R^2 was ever intended to be used to structure the data base, but it has become a statistical "workhorse" in regression analysis and it is important to note its limitation. Mahalanobis distance, first introduced in 1930 (Ref. 9), is a measure of analogy that could be used to compliment R^2 in deriving a CER which might be a better predictor of costs. Professor Wallenius has recently reintroduced Mahalanobis distance (Ref. 18) in this regard, and has created enough interest to attempt to determine its worth. It is discussed in Section V of this thesis.

The second basic concept contributing to the problem with existing CERs is the tendency to use them for applications other than those for which they were intended. Each situation for which an analyst chooses to use a CER, either as a primary or a back-up estimate, is unique with respect to what is required of the CER. The requirements may simply dictate that the best CER is the one that will provide an estimate the quickest, or these requirements may demand more of the CER.

When proposed system requirements are only tentative, the analyst's only concern is trade-offs among important decision variables, or

comparisons of alternative designs. A CER developed on a total cost basis with readily quantifiable explanatory variables, such as system performance characteristics, would be sufficient. The absolute accuracy of the CER would not be important as long as the relative accuracy is consistent and sensitive to the variables being traded-off. In other words, if the CER consistently over-estimated, or consistently under-estimated costs, it would still be of use to the analyst because it is the differences in costs that are the primary concern in this situation.

For evaluation of contractor proposals, a CER for each of the major cost accounts would be necessary. Absolute accuracy of the estimate would become more important, and explanatory variables that reflected such factors as contractor experience or maximum tooling capacity might be more appropriate.

It is apparent from all this that one model based on a limited number of CERs derived from the same data base, with perhaps some optional CERs or explanatory variables, probably is not going to be adequate to meet the demands of today's analyst.

To enhance the future use and benefits of CERs, the analyst must consider these two basic concepts before developing new models or improving upon existing ones. What is required is a set of guidelines by which the analyst may develop a CER for his specific purpose as a function of the type of cost estimate he desires and the characteristics of the airframe in question. Consideration should be given also to Mahalanobis distance as a means of determining the data base that is more apt to reflect performance characteristics similar to the proposed system.

IV. CONSIDERATIONS FOR THE FUTURE APPLICATION OF AIRCRAFT AIRFRAME CERs

A strategy to improve future independent parametric cost estimates would be to develop CERs for each specific proposed system for which the cost is to be estimated. In this way, optimal use of available information can be made by choosing candidates for the data base according to their analogy with the proposed system, and selecting among explanatory variables according to the nature of the costs and the ability to quantify them. To minimize the effort and to increase the effectiveness of this task with respect to aircraft airframe costs, it is important to draw upon previous experience. The data base and the explanatory variables are two aspects with which the analyst must be familiar.

The data base must include both cost and performance characteristics information. An accurate data base is the most important aspect in developing a meaningful CER. As discussed in Chapter I, the Rand Corporation has contributed significantly to collecting and "cleaning" the data base for aircraft airframe costs. This cleaning process entails many considerations. Despite the emphasis placed on uniform data collection by the Contractor Cost Data Reporting program, information is still received in varying formats. This is especially true when the data base spans many years.

The information collected has to be matched to the particular aircraft and the specific stage of production. A learning curve technique is used to adjust for differences in cost due to varying production quantities. Learning curve slopes can be calculated from the data if sufficient information exists, or estimates of previously

experienced learning curve slopes can be utilized. Cost for various quantities can then be estimated. Another aspect of this "matching" problem concerns derivative or prototype aircraft. The derivative aircraft generally will have gained some cost savings advantages because of the many similarities with the earlier production version. If these cost differences cannot be quantified, or the proposed system is of a derivative nature, it may not be appropriate to use a prototype design in the data base.

Definitional differences must be considered in cleaning the data. Cost categories are the obvious area where this occurs, but the definition of performance characteristics will cause inconsistencies also in the information. For example, gross take-off weight is a function of the amount of avionics installed, type and amount of armament, and fuel load. This results in different values of gross weight depending upon the mission requirements for which it is defined.

Adjustments for time also are required. Tooling, material, support, and other cost categories must be measured in dollars which vary through the years if for no other reason than inflation. Price indicies are used to correct for this problem; however, errors in the indicies themselves are introduced so their use should be limited. Ideally, those items that can be measured in hours should be left in hours to avoid having to correct for dollar value variation.

One final comment concerning cleaning the data is the effect on cost of different service imposed requirements for the same aircraft. The landing gear on Navy procured aircraft will include additional costs to strengthen them for carrier landings. This effect should be isolated and removed, or explained by the regression using a dummy variable.

This is by no means a conclusive discussion of the problems of data adjustments, nor is it intended to be. It is presented so that the analyst is aware of the implications in selecting candidates for the data base. Also, it should be recognized that this problem of establishing a reliable data base is a continuous one. It never can be resolved to complete satisfaction because of the dynamic nature of the environment.

Given a data base, the choice among explanatory variables is the second most important aspect in developing a reliable CER. There are many explanatory variables for which it can be argued that there is a causal relationship between their value and airframe costs. This results in an even larger number of possible combinations of explanatory variables that could be used in a regression equation. To consider all possible combinations is unnecessary. If two or more explanatory variables have similar effects on measuring variability in cost they are said to be correlated. Nothing is gained by including an additional explanatory variable that is highly correlated with a variable already present in the regression equation. If multicollinearity exists, then there is the added problem of interpreting coefficient values, as noted earlier.

To assist in minimizing the amount of correlation, explanatory variables may be grouped into functional categories. In determining a CER, normally the selection of explanatory variables would be limited to no more than one variable per functional category, and often there is even strong correlation between functional categories. The number of categories to include would depend upon the purpose for which the CER is intended.

Table Three is a summary of the more commonly used variables listed according to seven (7) functional categories. These categories include: Size, Military Usefulness, Construction, Range, Program Characteristics, and Maneuverability.

TABLE THREE

CATEGORIZED LIST OF EXPLANATORY VARIABLES*

(Compiled from Refs. 7, 8, & 15)

Size

Weight

Wetted Area

Wing Area

Construction/Design

Wing Type

Structural Efficiency Factor

Ratio of $\frac{\text{Total Weight--Airframe Weight}}{\text{Airframe Weight}}$

Skin Friction Drag

Max Lift Coefficient

Design Ultimate Load Factor

Carrier Capability

Program Characteristics

Contractor Experience

Tooling Capability

# of Test Aircraft

Index of Program Difficulty

New Engine Dummy Variable

Military Usefulness/Combat

Maximum Sustained Speed Capability

Maximum Climb Rate

Speed

Specific Power

Maximum Specific Energy

Range

Internal Fuel Fraction

Breguet Range Factor

Payload Fraction

Total Fuel Fraction

Maneuverability

Maximum Sustained Load Factor

Thrust to Weight Ratio

Wing Loading

Other

Objective Technology Index

Time

*See Appendix A for definition

From a simplistic point of view, size would be expected to affect cost in the sense that the more you have of something, the more it will cost. Use of an explanatory variable in this category is appropriate for many different CERs, but since it is highly correlated with others, it may be omitted from performance oriented applications. Military worth, range, and maneuverability could be considered as one functional category entitled "performance," but to do so would suppress important descriptive information. These performance related categories are especially useful early in the acquisition process because they are reasonably quantifiable, and the mission needs of a particular aircraft are normally addressed in these terms. Construction/Design oriented explanatory variables are used to account for differences in such things as structural strength, complexity of different wing configurations, fabrication technology, integration of avionics, and the like. Their use would be considered more appropriate as the proposed system becomes more defined.

Unfortunately, the size, performance and construction characteristics of airframes cannot explain all the variability in costs. Many costs are program related. They include contractor experience, tooling capability, availability of labor, number of test aircraft, advancement in the state of the art, capacity, and the like. These factors are not as quantifiable as other characteristics, and not all can be accounted for in a CER. The data base includes a wide assortment of programs. Therefore the CER will not be sensitive to small changes. Additionally, there is the implicit assumption that every program will have its fair share of technical, programming, and funding problems. To the extent that program related explanatory variables can be used, their application

is limited to the later stages of the acquisition process beginning with receipt and evaluation of contractor proposals.

V. MAHALANOBIS DISTANCE OR A MEASURE OF ANALOGY

Given a system whose cost is to be estimated, a data base of similar systems and a methodology for deriving a CER, there remains two key decisions in the development of a "good" CER: the choice among systems to be used in the data base, and the choice among various explanatory variables. These two decisions normally are treated as being independent.

The data base is specified first and usually includes all similar systems for which cost information is available. This was the case for the three (3) aircraft airframe models described in Section II. Some attempts have been made to stratify the sample so that the data base might reflect the proposed system better. One such stratification was according to aircraft type (e.g., fighter aircraft) and is detailed in Ref. 4. It was found that the fighter aircraft sample CERs were of poorer statistical quality and did not estimate costs for the four (4) most recent fighters in the data base as well as the total sample derived CERs.

Another attempt at stratifying the data base was by speed ranges. In both cases, the decision concerning stratification was made without considering the explanatory variables that would be used. Also, the stratification decision was not made relative to a specific proposed system, but rather to a category of systems in which a proposed system might be classified.

Both the choice of data base systems and the choice of explanatory variables are often made without considering the proposed system. This approach does not seem reasonable in light of the fact that the purpose of the CER is to estimate the cost of this system. It further supports

the contention in Section II of this thesis that CERs should be tailored to a specific system. Additionally, it is not apparent that these decisions should be made independently. If the data base is to be determined according to the relationship between values of explanatory variables of systems in the data base and the corresponding values of explanatory variables of the proposed system, it stands to reason that a choice of different explanatory variables could affect what systems would be most appropriate to include in the data base.

For example, if the proposed system is the F-4 and speed is to be used as an explanatory variable, the choice of historical aircraft is limited. All other previously manufactured aircraft have lower speeds, and only six (6) have speed capabilities reasonably comparable to the F-4. On the other hand, if wing area is considered as an explanatory variable, a range of values about the wing area of the F-4 exists, and there are ten (10) aircraft with wing area values comparable to the F-4 wing area.

A measure of this relationship between explanatory variable values of the data base and those of the proposed system is part of the calculation of prediction intervals and takes the form of $E' \ (X'X)^{-1} \ E$ (see Section III). Another related approach that has been introduced as a means of quantifying this relationship or analogy between the data base and the proposed system explanatory variables is Mahalanobis distance (MD). The formula for Mahalanobis Distance is: $MD = (\tilde{x} - \bar{x})' \ S^{-1} \ (\tilde{x} - \bar{x})$, where,

$\tilde{x}$ = the vector of the proposed system explanatory variable values

$\bar{x}$ = the vector of the data base system explanatory variable mean values

S = the covariance matrix of the data base system explanatory variable values.

The formula for the S matrix can be written in several ways, one of which is: $S = \frac{xx' - n\bar{x}\bar{x}'}{n - 1}$, where,

x = matrix of explanatory variable coefficients

n = number of systems in the data base

In this form, the relationship between MD and the $E' (XX')^{-1} E$ term of the prediction interval formula of Section III can be observed. Mahalanobis distance is a function of both the choice of explanatory variables and the systems in the data base. It is a measure of analogy in that the difference between the proposed system and data base system explanatory variable mean values are "weighted" by the S matrix. From the expression $(\tilde{x} - \bar{x})$ it is clear that the closer the proposed system values are to the data base mean values, the smaller the Mahalanobis distance becomes, and therefore, the greater is the analogy between data base and proposed system.

The effects on MD caused by variation in S is not clear, but must be understood if the analyst is to use MD as a means of improving the analogy of the data base and the proposed system. An alternative formula for the elements of the S matrix is: $S_{ij} = \frac{1}{n-1} \sum_{h=1}^{n} (x_{hi} - \bar{x}_i)(x_{hj} - \bar{x}_j)$

where,

n = number of explanatory variables

k = number of explanatory variables

x = $n \times k$ matrix, each column of which contains the values of an explanatory variable for each system in the data base.

S will be a $k \times k$ symetric matrix whose diagonal elements will be the variance of the ith explanatory variable (σ_i ; $i = 1, 2, \cdots, k$) and whose off-diagonal elements will be the covariance between explanatory variables.

Assuming for the moment that the covariance between explanatory variables would be zero (0), the S matrix would take the following form:

$$\begin{bmatrix} \sigma_1^2 & & & & \\ & \sigma_2^2 & & & \\ & & \cdot & & \\ & & & \cdot & \\ & & & & \cdot \\ & & & & & \sigma_k^2 \end{bmatrix}$$

(all other elements would be 0)

It is easy to show from $(XX^{-1}) = I$ that the inverse of this matrix would be:

$$\begin{bmatrix} 1/\sigma_1^2 & & & & \\ & 1/\sigma_2^2 & & & \\ & & \cdot & & \\ & & & \cdot & \\ & & & & \cdot \\ & & & & & 1/\sigma_k^2 \end{bmatrix}$$

and therefore, the calculation of MD would reduce to: $MD = \sum_{j=1}^{k} \frac{(x_j - \bar{x}_j)^2}{\sigma_j^2}$

where: k, x, and $\bar{x}$ are defined as before.

In this form, which assumes no covariance between explanatory variables, it can be seen that increases in variability (σ_j^2) of the $_j$th data base system explanatory variable will reduce MD. The immediate implication of this is that it is not optimal simply to choose data base systems whose explanatory variable values compare closely to the proposed system values. The optimal approach is to introduce as much variability as possible while maintaining a mean value close to the proposed system value. There is an intuitive side to this in the sense that the greater the dispersion between two points the more confidence one has in fitting a line between them.

The reasonableness of the assumption that the covariance is zero (0) must be considered. The covariance and correlation between two

explanatory variables are related by the following expression:

$$\rho_{xy} = \frac{\text{covariance } (x,y)}{\sigma_x \sigma_y} \qquad \text{where,}$$

ρ = the correlation coefficient

x and y are two arbitrary explanatory variables with variances σ_x and σ_y. Obviously there will be no correlation between explanatory variables only when the covariance between explanatory variables is zero (0).

In developing a CER it has been noted that the correlation between explanatory variables should be minimized in order to avoid sporadic results implying that the assumption of zero (0) or minimum covariance is reasonable. However, regardless of the desire to minimize correlation, it will always exist to some extent, and therefore its effects, along with the effects of variability on Mahalanobis distance should be examined.

The effect of variability on MD can be demonstrated by considering the following matrix which represents hypothetical values of three (3) different explanatory variables (columns) and four (4) systems in the data base (rows). The assumption of zero (0) covariance will no longer hold, but if it is kept reasonably constant the effects of variability should be observed.

$$A = \begin{bmatrix} 4 & 3 & 8 \\ 6 & 3 & 9 \\ 7 & 4 & 6 \\ 3 & 6 & 5 \end{bmatrix}$$

where: column variances are 3.3, 2, and 3.3

column means are 5, 4, and 7

For a proposed system whose corresponding explanatory variable values are 7, 6, and 8: MD = 41.10

By introducing some more variability into the values of the first explanatory variable while holding the mean constant, the A matrix becomes:

$$A_1 = \begin{bmatrix} 1 & 3 & 8 \\ 9 & 3 & 9 \\ 6 & 4 & 6 \\ 4 & 6 & 5 \end{bmatrix}$$

where: column variances are 11.3, 2, and 3.3

column means are 5, 4, and 7

For the same proposed system, MD = 20.67. The increase in variability of just one of the explanatory variables has reduced MD.

Repeating the process by introducing more variability into the values of the second explanatory variable, the A_1 matrix becomes:

$$\begin{bmatrix} 1 & 1 & 8 \\ 9 & 4 & 9 \\ 6 & 10 & 6 \\ 4 & 1 & 5 \end{bmatrix}$$

where: column variances are 11.3, 18, and 3.3

column means are 5, 4, and 7

For the same proposed system MD = .64. Again, by increasing the variance of the explanatory variables the Mahalanobis distance has been reduced. By examining the complete covariance matricies (CVA, CVA_1, CVA_2) of the three example matricies (A, A_1, A_2) an understanding of the potential effects of covariance on MD can be observed.

$$CVA = \begin{bmatrix} 3.3 & -1.3 & 1 \\ -1.3 & 2 & -2.3 \\ 1 & -2.3 & 3.3 \end{bmatrix} \quad CVA_1 = \begin{bmatrix} 11.3 & -.67 & 1.67 \\ -.67 & 2 & -2.3 \\ 1.67 & -2.3 & 3.3 \end{bmatrix} \quad CVA_2 = \begin{bmatrix} 11.3 & 7 & 1.67 \\ 7 & 18 & -1 \\ 1.67 & -1 & 3.3 \end{bmatrix}$$

The covariances remained relatively constant as more variability was introduced, with the possible exception of the covariance between the first and second explanatory variables in CVA_2 which increased from -0.67 to 7.

To illustrate potential effects of covariance on MD, more variability was introduced into the values of the third explanatory variable while simultaneously trying to establish more correlation between variables. The A_2 and CVA_2 matricies became:

$$A_3 = \begin{bmatrix} 1 & 1 & 1 \\ 9 & 4 & 9 \\ 6 & 10 & 15 \\ 4 & 1 & 3 \end{bmatrix} \qquad CVA_3 = \begin{bmatrix} 11.3 & 7 & 14.67 \\ 7 & 18 & 26 \\ 14.67 & 26 & 40 \end{bmatrix}$$

For the same proposed system MD = 187.23

The variance of the third explanatory variable was substantially increased from 3.3 to 40, but the expected reduction in MD was more than

offset by increases in the covariance (1.67 to 14.67 between the first and third variables, and -1 to 16 between the second and third variables). The off-diagonal elements of CVA_3 are large compared to the diagonal elements which was not the case for CVA, CVA_1, and CVA_2. The obvious implication is that increases in covariance increase the Mahalanobis distance.

Taking this example one step further, the variances of the explanatory variables were fixed, as are the mean values, but the covariances were reduced by changing the order of elements within columns. The A_3 and CVA_3 matrices became:

$$A_4 = \begin{bmatrix} 1 & 4 & 9 \\ 9 & 1 & 1 \\ 6 & 1 & 15 \\ 4 & 10 & 3 \end{bmatrix} \qquad CVA_4 = \begin{bmatrix} 11.3 & -7 & -6.67 \\ -7 & 18 & -10 \\ -6.67 & -10 & 40 \end{bmatrix}$$

For the same proposed system MD = 2.53

The reduction in covariance had the anticipated effect of reducing MD. It is apparent that if the object is to minimize MD, then the choice among explanatory variables should be such that the covariance is minimized. This effect of covariance on MD tends to support the notion introduced earlier of minimizing collinearity in the choice among data base systems and explanatory variables.

This is by no means a complete examination of the effects of variability and covariance on MD. For example, the signs of the covariance elements if mixed could have offsetting effects causing large covariance to go unnoticed. However, it must be remembered that the overriding considerations when choosing among data base systems and explanatory variables is an understanding of the system and the causal relationships that exist. Mahalanobis distance, as discussed here, is only a means

of assisting the analyst in achieving a more reliable CER by dealing with the issue of analogy between the data base and the proposed system.

VI. SUMMARY

There is a recognized need for the use of independent parametric cost estimates in the acquisition of major weapons systems. Through the years, considerable effort has been expended in deriving reliable cost estimating relationships (CERs) to fulfill this need. To date, the majority of models developed are applicable to "types" of systems rather than to a specific system. In particular, the models developed for aircraft airframe costs are applicable to any reasonably similar future aircraft airframe which might be proposed. This approach seems unreasonable in the sense that the CER will be applied to a specific proposed airframe, yet the CER is developed when little or nothing is known about the characteristics of this proposed airframe.

A strategy to improve future independent parametric cost estimates would be to develop CERs for a specific proposed system. In this way, optimal use of available information can be made, and consideration can be given to the analogy with the proposed system for various choices of data base systems and explanatory variables.

This approach is feasible only if the analyst draws upon previous experience in CER development. Two areas are important in this regard. The analyst must have a current data base and must be familiar with any adjustments that were made due to inconsistencies in the information and inconsistencies that might still remain. Additionally, the choice of explanatory variables should be guided by previous experience concerning both the causal relationships that have existed with cost and the problems with multicollinearity that have occurred.

Mahalanobis distance (MD) has been introduced as a means to assist the analyst in choosing a combination of data base systems and explanatory variables that will be more analogous to the proposed system thereby resulting in a potentially more reliable CER. It has been shown, in general, that MD can be minimized by reducing collinearity and increasing variability among data base performance characteristics while attempting to maintain the mean values of these performance characteristics "close" to the corresponding values of the proposed system performance characteristics.

APPENDIX A

DEFINITIONS OF SELECTED EXPLANATORY VARIABLES

Breguet Range Factor: The product of cruise speed and lift-to-drag ratio divided by the specific fuel consumption.

Combat Weight: Weight of an aircraft with full internal ordnance and 60% of its internal fuel capacity remaining.

Design Ultimate Load Factor: The maximum load factor the aircraft is designed to withstand at the stress design weight without structural failure.

Internal Fuel Fraction: Weight of internal fuel capacity divided by the difference between full internal weight and weight of internal fuel capacity.

Maximum Specific Energy: The maximum sum of kinetic and potential energy developed at 1 G level flight divided by combat weight.

Maximum Sustained Speed Capability: Maximum speed of an aircraft at combat weight.

Payload Fraction: The difference between gross weight and internal weight divided by gross weight.

Specific Power: The product of maximum static thrust and maximum velocity divided by combat weight.

Structural Efficiency Factor: The structure weight divided by the product of design stress weight and ultimate load factor.

Sustained Load Factor: Maximum load factor the aircraft can sustain in level flight at combat weight at an altitude of 25,000 feet and a Mach number of 0.8.

Wetted Area: Total surface area of the aircraft.

Wing Loading: Combat weight divided by wing area.

(compiled from Refs. 7 and 15)

LIST OF REFERENCES

1. Baker, F. N., Improving the Cost Estimating and Analysis in DOD and NASA, PH.D. Thesis, George Washington University, Washington, D.C., 1972.

2. Boren, H. E., A Computer Model for Estimating Development and Procurement Costs of Aircraft (DAPCA--III), Rand Corporation Report R-1884-PR, March 1976.

3. Draper, N. R. and Smith, H., Applied regression Analysis, New York: John Wiley and Sons, Inc., 1966.

4. Fisher, G. H., Cost Considerations in Systems Analysis, New York: American Elsevier, 1971, Chapters 6, 7, and 8.

5. Hess, R. W., Fighter Aircraft Airframe Cost-Estimating Relationships: An Exploratory Analysis, Rand Corporation Working Note RAND/WN-10180-AF, Privileged Information, May 1978.

6. Large, J. P. and Balchelder, L. J., Aircraft Airframe Cost Data, Rand Corporation Working Note RAND/WN-10227/1-PA & E, Privileged Information, August 1978.

7. Large, J. P., Campbell, H. G., and Cater, D., Parametric Equations for Estimating Aircraft Airframe Costs, Rand Corporation Report R-1693-PA & E, May 1975.

8. Large, J. P. and Gillespie, Capt. K. M. S., A Critique of Aircraft Airframe Cost Models, Rand Corporation Report R-2194-AF, September 1977.

9. Mahalanobis, P. C., On Tests and Measures of Group Divergence, J. Asiat. Soc. Beng., (26), 1930.

10. Maiolo, B. A., The Use of Parametric Cost Models in Industry, paper presented at the 31st MORS Working Session B-4, U.S. Naval Academy, Annapolis, Maryland, 19-21 June 1973.

11. Miller, B. M. and Sovereign, M. G., Parametric Cost Estimating with Applications to Sonar Technology, Naval Postgraduate School NPS 55Z073091A, September 1973.

12. Noah, J. W. and others, Estimating Aircraft Acquisition Costs by Parametric Methods, J. Watson Noah Associates Report FR-103-USN, September 1973.

13. Aircraft Cost Handbook, TR-103-USN, J. Watson Noah Associates, Inc., Privileged Information, July 1977.

14. Pindyck, R. S. and Rubinfeld, D. L., Econometric Models and Economic Forecasts, McGraw-Hill, Inc., 1976.

15. Stanley, W. L and Miller, M. D., Measuring Technological Change in Jet Fighter Aircraft, Rand Corporation Report R-2249-AF, September 1972.

16. Theil, H., Principles of Econometrics, Wiley, 1971.

17. Timson, F. S. and Tihansky, D. P., Confidence in Estimated Airframe Costs: Uncertainty Assessment in Aggregate Predictions, Rand Corporation Report R-903-PR, October 1972.

18. Wallenius, K. T., Mahalanobis Distance as a Measure of Analogy in Parametric Cost Estimation, Clemson University Report No. N98, Technical Report #322, September 1978.

INITIAL DISTRIBUTION LIST

		No. Copies
1.	Defense Technical Information Center Cameron Station Alexandria, Virginia 22314	2
2.	Library, Code 0142 Naval Postgraduate School Monterey, California 93940	2
3.	Professor Michael G. Sovereign, Code 55 Chairman, Department of Operations Research Naval Postgraduate School Monterey, California 93940	10
4.	Professor Donald P. Gaver, Code 55 Gv Department of Operations Research Naval Postgraduate School Monterey, California 93940	1
5.	LCDR Bruce Robert Bennett, USN 1529 Apaloosa Court Carson City, Nevada 89701	1

Thesis
B3693 Bennett 189338
c.1
The use of parametric cost estimating relationships as they pertain to aircraft airframes; a new perspective.